Yibo Zhu

ByteDance Inc.
601 108th Ave NE, Ste 1580
Bellevue, WA 98004

bobzhuyb@gmail.com
http://yibozhu.com
(805) 886-2536

Work Experience

ByteDance Inc., Bellevue, WA 10/18 – present
Director of Machine Learning Systems, Applied Machine Learning

I am currently a director managing 100+ engineers (30 in the US) and overseeing the R&D of machine learning infrastructure at ByteDance. I joined as a (staff-level) Research Scientist and started the first distributed GPU training system of ByteDance. Later, my scope grew to include inference systems, ML platforms, resource scheduling, etc. Selected projects and business impact:

- Our infrastructure started with <1K GPUs and expanded to **>100K** GPUs now. *All* ByteDance's neural network models run on our infrastructure, including recommendation, search, content moderation, video visual effects, translation, audio, game AI, etc.
- We provide ML platforms that are used by **thousands of** machine learning engineers, from flexible web IDE and low-level debugging environment for expert users, to convenient codeless model development for non-expert users.
- We designed cluster scheduling strategies that increased GPU *allocation* rate for training from 70% to **95%+**; increased the average GPU *utilization* for both training and inference from ~30% to **~70%**. Saved hundreds of millions of dollars of GPU costs.
- I started and was one of the first developers of BytePS (https://github.com/bytedance/byteps), an open-sourced high-performance parameter server framework. An internal version is the cornerstone of distributed training at ByteDance.
- I proposed, designed and recruited the team for ByteDance's internal ML compiler stack. It compiles production TensorFlow and PyTorch models onto GPU and other AI chips with optimizations. Open-source version: https://github.com/bytedance/byteir
- My team started the I/O and data preprocessing pipeline dedicated to DNN workloads. It is now widely used in ByteDance. Open sourced at https://github.com/bytedance/matxscript and https://github.com/CVCUDA/CV-CUDA
- In addition, my team and I started and own ByteDance's niche ML frameworks for: reinforcement learning, giant multi-modality models, sparse mixture-of-expert models, elastic AutoML training, graph neural networks on GPU, etc.

Microsoft Corp., Redmond, WA 8/16 – 10/18
Senior Researcher II, Mobility and Networking Research Group

I worked on large-scale network and systems for Microsoft Azure. I led the following projects that were deployed and used by Azure.

- DCQCN: the de facto standard RDMA congestion control protocol in the industry, adopted by all Mellanox NICs and Azure.
- CrystalNet: cloud-scale network emulator using containers and customized virtual networking.
- Everflow: Large-scale network telemetry for network troubleshooting.
- Freeflow: enabled RDMA and accelerated TCP/IP stack over container overlay network, for running applications with Kubernetes. Open-sourced at GitHub: https://github.com/Microsoft/Freeflow

Education

University of California, Santa Barbara, Santa Barbara, CA 9/11 – 7/16
Ph.D., Computer Science
Advisor: Prof. Heather Zheng and Prof. Ben Y. Zhao

Tsinghua University, Beijing, China 8/07 – 7/11
Bachelor of Electronic Engineering (GPA top 1%, graduated with distinction among all college students in Beijing)

Stanford University, Stanford, CA 7/10 – 9/10
Exchange student, Undergraduate Visit and Research program

Pre-graduation Industrial Experience

Microsoft Research, Redmond, WA 6/13 – 9/15
Research Intern, Mobility and Networking Research Group

- Multi-year internship mentored by Dr. Jitu Padhye, Dr. Ming Zhang and Dr. Ratul Mahajan

- Multiple networking projects, including the first RDMA deployment at Azure, network diagnosis systems and the founding phase of SONiC (https://github.com/Azure/SONiC)

Microsoft Research Asia, Beijing, China 9/10 – 3/11
Research Intern, Wireless and Networking Group

- Worked with Dr. Chuanxiong Guo. Designed and implemented a reliable group data delivery system for datacenters.

Academic Experience

UCSB SAND Laboratory, Santa Barbara, CA 9/11 – 6/16

- Designed new wireless primitives for augmenting bandwidth and building facilities networks in datacenters.
- Explored the feasibility of using 60GHz in cellular network and mobile sensing for orders of magnitude performance gain over traditional WiFi/LTE-based approaches.
- Measured and analyzed malicious crowdsourcing systems targeting today's online social networks.

Tsinghua University NGN Laboratory, Beijing, China 2/09 – 7/11

- Worked with Prof. Xing Li and Prof. Beixing Deng. Designed and implemented *Toread*, a decentralized network-coordinate system on *PlanetLab*. Project homepage: http://code.google.com/p/toread

Awards

MSR Redmond Labs Exemplary Collaboration Award (2017): awarded to the best technology transfer.
Microsoft Research Fellowship (2015): annually awarded to 12 Ph.D. students in North America.
UCSB Holbrook Fellowship (2011): annually awarded to 6 Ph.D. freshmen in UCSB.
Student Travel Grant, ICNP'10, NSDI'12, DySPAN'12, HotMobile'15

Pre-graduate school awards:
Graduated with distinction among all college students in Beijing city (2011)
Chinese National Scholarship (2008-2010): top 3% students of Tsinghua University
1st Place, Programming Competition in Department of EE, Tsinghua University (2008)
Golden Medal in 22th Chinese Mathematical Olympiad (2007): top 30 of mainland China

Publications

[1] Yangrui Chen, Cong Xie, Meng Ma, Juncheng Gu, Yanghua Peng, Haibin Lin, Chuan Wu, **Yibo Zhu**, SAPipe: Staleness-Aware Pipeline for Data Parallel DNN Training. *To appear in NeuIPS'22.*

[2] Zhuang Wang, Haibin Lin, **Yibo Zhu**, T.S. Eugene Ng, Espresso: Revisiting Gradient Compression from the System Perspective. *To appear in EuroSys'23.*

[3] Tianfeng Liu, Yangrui Chen, Dan Li, Chuan Wu, **Yibo Zhu**, Jun He, Yanghua Peng, Hongzheng Chen, Hongzhi Chen, Chuanxiong Guo, BGL: GPU-Efficient GNN Training by Optimizing Graph Data I/O and Preprocessing. *To appear in NSDI'23.*

[4] Yihao Zhao, Yuanqiang Liu, Yanghua Peng, **Yibo Zhu**, Xuanzhe Liu, Xin Jin, Multi-Resource Interleaving for Deep Learning Training. *In Proc. of SIGCOMM'22.*

[5] Hanpeng Hu, Chenyu Jiang, Yuchen Zhong, Yanghua Peng, Chuan Wu, **Yibo Zhu**, Haibin Lin, Chuanxiong Guo, dPRO: A Generic Performance Diagnosis and Optimization Toolkit for Expediting Distributed DNN Training. *In Proc. of MLSys'22.*

[6] Jiarong Xing, Leyuan Wang, Shang Zhang, Jack Chen, Ang Chen, **Yibo Zhu**, Bolt: Bridging the Gap between Auto-tuners and Hardware-native Performance. *In Proc. of MLSys'22.*

[7] Xinhao Kong, **Yibo Zhu**, Huaping Zhou, Zhuo Jiang, Jianxi Ye, Chuanxiong Guo, Danyang Zhuo, Collie: Finding Performance Anomalies in RDMA Subsystems, *In Proc. of NSDI'22.*

[8] Yuchen Jin, Tianyi Zhou, Liangyu Zhao, **Yibo Zhu**, Chuanxiong Guo, Marco Canini, Arvind Krishnamurthy, AutoLRS: Automatic Learning-Rate Schedule by Bayesian Optimization on the Fly. *In Proc. of ICLR'21.*

[9] Hwijoon Lim, Wei Bai, **Yibo Zhu**, Youngmok Jung, Dongsu Han, Towards Timeout-less Transport in Commodity Datacenter Networks. *In Proc. of EuroSys'21.*

[10] Yimin Jiang, **Yibo Zhu,** Chang Lan, Bairen Yi, Yong Cui, Chuanxiong Guo, A Unified Architecture for Accelerating Distributed DNN Training in Heterogeneous GPU/CPU Clusters. *In Proc. of OSDI'20.*

[11] Zhihao Bai, Zhen Zhang, **Yibo Zhu,** Xin Jin, PipeSwitch: Fast Pipelined Context Switching for Deep Learning Applications. *In Proc. of OSDI'20.*

[12] Daehyeok Kim, Zaoxing Liu, **Yibo Zhu**, Changhoon Kim, Jeongkeun Lee, Vyas Sekar, Srinivasan Seshan, TEA: Enabling State-Intensive Network Functions on Programmable Switches. *In Proc. of SIGCOMM'20.*

[13] Yangrui Chen, Yanghua Peng, Yixin Bao, Chuan Wu, **Yibo Zhu,** Chuanxiong Guo, Elastic Parameter Server Load Distribution in Deep Learning Clusters. *In Prof. of SoCC'20.*

[14] Yanghua Peng, **Yibo Zhu,** Yangrui Chen, Yixin Bao, Bairen Yi, Chang Lan, Chuan Wu, Chuanxiong Guo, A Generic Communication Scheduler for Distributed DNN Training Acceleration. *In Proc. of SOSP'19.*

[15] Danyang Zhuo, Kaiyuan Zhang, **Yibo Zhu,** Hongqiang Harry Liu, Matthew Rockett, Arvind Krishnamurthy, Thomas Anderson, Slim: OS Kernel Support for a Low-Overhead Container Overlay Network. *In Proc. of NSDI'19.*

[16] Da Yu, **Yibo Zhu,** Behnaz Arzani, Rodrigo Fonseca, Tianrong Zhang, Lihua Yuan, Karl Deng, dShark: A General, Easy to Program and Scalable Framework for Analyzing In-network Packet Traces. *In Proc. of NSDI'19.*

[17] Daehyeok Kim, Tianlong Yu, Hongqiang Harry Liu, **Yibo Zhu**, Jitu Padhye, Shachar Raindel, Chuanxiong Guo, Vyas Sekar, Srinivasan Seshan, FreeFlow: Software-based RDMA Virtual Networking for Containerized Clouds. *In Proc. of NSDI'19.*

[18] Juncheng Gu, Mosharaf Chowdhury, Kang G. Shin, **Yibo Zhu,** Myeongjae Jeon, Junjie Qian, Hongqiang Liu, Chuanxiong Guo, Tiresias: A GPU Cluster Manager for Distributed Deep Learning. *In Proc. of NSDI'19.*

[19] Gaoxiong Zeng, Wei Bai, Ge Chen, Kai Chen, Dongsu Han, **Yibo Zhu,** Lei Cui, Congestion Control for Cross-Datacenter Networks. *In Proc. of ICNP'19.*

[20] Daehyeok Kim, Amirsaman Memaripour, Anirudh Badam, **Yibo Zhu**, Hongqiang Harry Liu, Jitu Padhye, Shachar Raindel, Vyas Sekar, Srinivasan Seshan, Steven Swanson, HyperLoop: Group-Based NIC-Offloading to Accelerate Replicated Transactions in Multi-Tenant Storage Systems. *In Proc. of ACM SIGCOMM'18.*

[21] Behnaz Arzani, Selim Ciraci, Luiz Chamon, **Yibo Zhu**, Hongqiang Liu, Jitendra Padhye, Geoff Outhred, Boon Thau Loo, Democratically Finding the Cause of Packet Drops. *In Proc. of USENIX NSDI'18*.

[22] Shuihai Hu, **Yibo Zhu**, Peng Cheng, Chuanxiong Guo, Kun Tan, Jitendra Padhye, Kai Chen, Tagger: Practical PFC Deadlock Prevention in Data Center Networks. *In Proc. of ACM CoNEXT'17*.

[23] Hongqiang Harry Liu*, **Yibo Zhu***, Jitu Padhye, Jiaxin Cao, Sri Tallapragada, Nuno P. Lopes, Andrey Rybalchenko, Guohan Lu, Lihua Yuan, CrystalNet: Faithfully Emulating Large Production Networks. *In Proc. of ACM SOSP'17*. *Co-primary authors

[24] Gaoxiong Zeng, Wei Bai, Ge Chen, Kai Chen, Dongsu Han, **Yibo Zhu**, Combining ECN and RTT for Datacenter Transport. *In Proc. of ACM APNet'17.*

[25] **Yibo Zhu**, Monia Ghobadi, Vishal Misra, Jitendra Padhye, ECN or Delay: Lessons Learnt from Analysis of DCQCN and TIMELY. *In Proc. of ACM CoNEXT'16.*

[26] Yanzi Zhu, **Yibo Zhu**, Ana Nika, Ben Y. Zhao, Haitao Zheng, Trimming the Smartphone Network Stack. *In Proc. of ACM HotNets'16.*

[27] Shuihai Hu, **Yibo Zhu**, Peng Cheng, Chuanxiong Guo, Kun Tan, Jitendra Padhye, Kai Chen, Deadlocks in Datacenter Networks: Why Do They Form, and How to Avoid Them. *In Proc. of ACM HotNets'16.*

[28] Ana Nika, Zhijing Li, Yanzi Zhu, **Yibo Zhu**, Ben Y. Zhao, Xia Zhou and Haitao Zheng, Empirical Validation of Commodity Spectrum Monitoring. *In Proc. of ACM SenSys'16.*

[29] Yanzi Zhu, **Yibo Zhu**, Ben Y. Zhao and Haitao Zheng, Reusing 60GHz Radios for Mobile Radar Imaging. *In Proc. of ACM MobiCom 2015.*

[30] **Yibo Zhu**, Daniel Firestone, Chuanxiong Guo, Jitendra Padhye, Shachar Raindel, Ming Zhang, Yehonatan Liron, Haggai Eran, Mohamad Haj Yahia and Marina Lipshteyn, Congestion Control for Large-scale RDMA Deployments. *In Proc. of ACM SIGCOMM 2015.*

[31] **Yibo Zhu**, Nanxi Kang, Jiaxin Cao, Albert Greenberg, Guohan Lu, Ratul Mahajan, Dave Maltz, Lihua Yuan, Ming Zhang, Haitao Zheng and Ben Zhao, Packet-Level Telemetry in Large Datacenter Networks. *In Proc. of ACM SIGCOMM 2015*.

[32] Ana Nika, **Yibo Zhu**, Ning Ding, Abhilash Jindal, Y. Charlie Hu, Xia Zhou, Ben Zhao and Haitao Zheng, Energy and Performance of Smartphone Radio Bundling in Outdoor Environments. *In Proc. of WWW 2015.*

[33] **Yibo Zhu**, Yanzi Zhu, Zengbin Zhang, Ben Y. Zhao and Haitao Zheng, 60GHz Mobile Imaging Radar. *In Proc. of ACM HotMobile 2015*.

[34] **Yibo Zhu**, Zengbin Zhang, Zhinus Marzi, Chris Nelson, Upamanyu Madhow, Ben Y. Zhao and Haitao Zheng, Demystifying 60GHz Outdoor Picocells. *In Proc. of ACM MobiCom 2014.*

[35] **Yibo Zhu**, Xia Zhou, Zengbin Zhang, Lin Zhou, Amin Vahdat, Ben Y. Zhao and Haitao Zheng, Cutting the Cord: A Robust Wireless Facilities Network for Data Centers. *In Proc. of ACM MobiCom 2014.*

[36] Jiaxin Cao, Chuanxiong Guo, Guohan Lu, Yongqiang Xiong, Yixin Zheng, Yongguang Zhang, **Yibo Zhu**, Chen Chen and Ye Tian, *Datacast:* A Scalable and Efficient Reliable Group Data Delivery Service for Data Centers. *In IEEE JSAC, 31(12):2632-2645, 2013.*

[37] Jiaxin Cao, Chuanxiong Guo, Guohan Lu, Yongqiang Xiong, Yixin Zheng, Yongguang Zhang, **Yibo Zhu** and Chen Chen, *Datacast:* A Scalable and Efficient Reliable Group Data Delivery Service for Data Centers. *In Proc. of ACM CoNEXT 2012.*

[38] Xia Zhou, Zengbin Zhang, **Yibo Zhu**, Yubo Li, Saipriya Kumar, Amin Vahdat, Haitao Zheng and Ben Y. Zhao, Mirror Mirror on the Ceiling: Flexible Wireless Links for Data Centers. *In Proc. of ACM SIGCOMM 2012.*

[39] Gang Wang, Christo Wilson, Xiaohan Zhao, **Yibo Zhu**, Manish Mohanlal, Haitao Zheng and Ben Y. Zhao, Serf and Turf: Crowdturfing for Fun and Profit. *In Proc. of WWW 2012.*

[40] **Yibo Zhu**, Yang Chen, Zengbin Zhang, Xiaoming Fu, Dan Li, Beixing Deng, Xing Li. Taming the Triangle Inequality Violations with Network Coordinate System on Real Internet. *In Proc. of ReArch'10 held in conjunction with CoNEXT'10.*

Talks

[1] Maximizing GPU utilization in Large Scale Machine Learning Infrastructure
[September 2022] *NVIDIA GTC, online*
[2] Understand and Leverage Heterogeneity in Machine Learning Clusters
[June 2022] *UC Berkeley, Berkeley, USA*
[3] HyperLoop: Group-Based NIC-Offloading to Accelerate Replicated Transactions in Multi-Tenant Storage Systems
[December 2018] *ChinaSys'18, Changsha, China*
[4] Novel Network Primitives Enabled by RDMA
[July 2018] *University of Wisconsin-Madison, Madison, USA*
[5] ECN or Delay: Lessons Learnt from Analysis of DCQCN and TIMELY
[December 2016] *CoNEXT'16, Irvine, USA*
[6] Network Design for the Cloud (job talk)
[March 2016] *Columbia University, New York City, USA*
[April 2016] *Microsoft Research, Redmond, USA*
[7] Congestion Control for Large-scale RDMA Deployments
[December 2015] *Google Networking Team, Mountain View, USA.*
[August 2015] *SIGCOMM'15, London, U.K.*
[September 2013] *Microsoft Azure Networking Team, Redmond, USA.*
[8] Packet-Level Telemetry in Large Datacenter Networks
[December 2015] *Google Networking Team, Mountain View, USA.*
[August 2015] *SIGCOMM'15, London, U.K.*
[9] 60GHz Mobile Imaging Radar
[February 2015] *HotMobile'15, Santa Fe, USA.*
[10] Cutting the Cord: A Robust Wireless Facilities Network for Data Centers
[September 2014] *MobiCom'14, Maui, USA.*
[11] Demystifying 60GHz Outdoor Picocells

[September 2014] *MobiCom'14, Maui, USA.*

[12] Taming the Triangle Inequality Violations with Network Coordinate System on Real Internet
[November 2010] *ReArch'10, held in conjunction with CoNEXT'10, Philadelphia, USA.*

Selected Press

[1] Microsoft reveals network simulator that keeps Azure alive. The Register, November 1, 2017.
[2] Microsoft's 'CrystalNet' Azure-network emulator may be available to customers one day. ZDNet, Octobor 31, 2017.
[3] Microsoft Azure Cloud Switch Is A Cross-Platform Linux-Based Operating System. Tech Times, September 20, 2015.
[4] Microsoft demonstrates its Linux-based Azure Cloud Switch operating system. ZDNet, September 18, 2015.
[5] Going wireless in the data center. ComputerWorld, May 7, 2012.
[6] Bouncing Data. MIT Technology Review, February 21, 2012.
[7] A Wireless Road Around Data Traffic Jams. New York Times, January 14, 2012.
[8] Speeding up the Internet by bouncing data off the ceiling. ExtremeTech, December 20, 2011.
[9] Million Dollar Crowdturfing Industry Dupes Social Networks, SlashDot, December 13, 2011.
[10] Hidden Industry Dupes Social Media Users, MIT Technology Review, December 12, 2011.

Professional Activities

[1] USENIX ATC'23, TPC, 2023
[2] USENIX ATC'22, TPC, 2022
[3] ACM SIGCOMM'21, TPC, 2021
[4] ACM SIGCOMM'20, TPC, 2020
[5] ACM HotNets'18, General chair, 2018
[6] ACM CoNEXT'18, TPC, 2018
[7] ACM SIGCOMM'18, TPC, 2018
[8] ACM SIGCOMM'18 KBNets Workshop, TPC, 2018
[9] ACM SIGCOMM'17 KBNets Workshop, Co-chair, 2017
[10] IEEE/ACM Transactions on Networking (ToN), Reviewer, 2016, 2017, 2018
[11] IEEE Transactions on Network and Service Management (TNSM), Reviewer, 2016
[12] IEEE Wireless Communications Letters, Reviewer, 2016
[13] Springer Journal of Network and Systems Management (JONS), Reviewer, 2016
[14] Transactions on Emerging Telecommunications Technologies (ETT), Reviewer, 2016
[15] MobiCom'15 S3 workshop, TPC, 2015.
[16] Elsevier Journal of Parallel and Distributed Computing (JPDC), Reviewer, 2015, 2016.
[17] IEEE Transactions on Mobile Computing (TMC), Reviewer, 2015.
[18] IEEE Transactions on Communications (TCOM), Reviewer, 2014, 2015.

Teaching

[1] CS276, Graduate Networking, *Grader,* UCSB, 2012.
[2] CS176B, Undergraduate Advanced Networking, *Teaching Assistant,* UCSB, 2012.
[3] CS176A, Undergraduate Networking, *Teaching Assistant*, UCSB, 2011.